GlideinWMS at the Holland Computing Center

Derek Weitzel

dweitzel@cse.unl.edu

HCC's usage of GlideinWMS

- Many users of GlideinWMS
  - Mathematics – Graph Existence
  - Computer Science – My Masters thesis
  - Bioinformatics – NMR (Rosetta, DaliLite)
  - Physics – LSST prototype
  - Digital Humanities – Every permutation ringtone.

- Exclusively use the GlideinWMS RPM
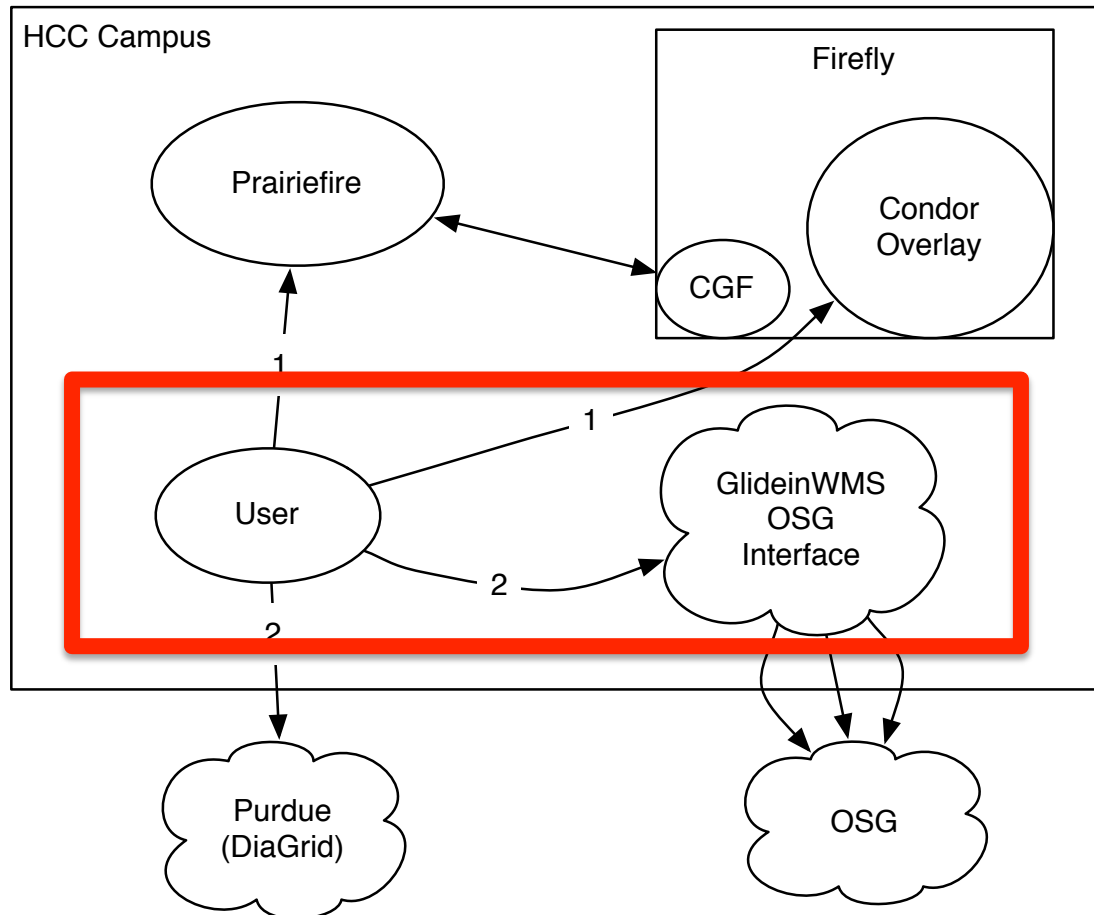 - Currently running glideinwms-vofrontend-2.5.4-2.osg

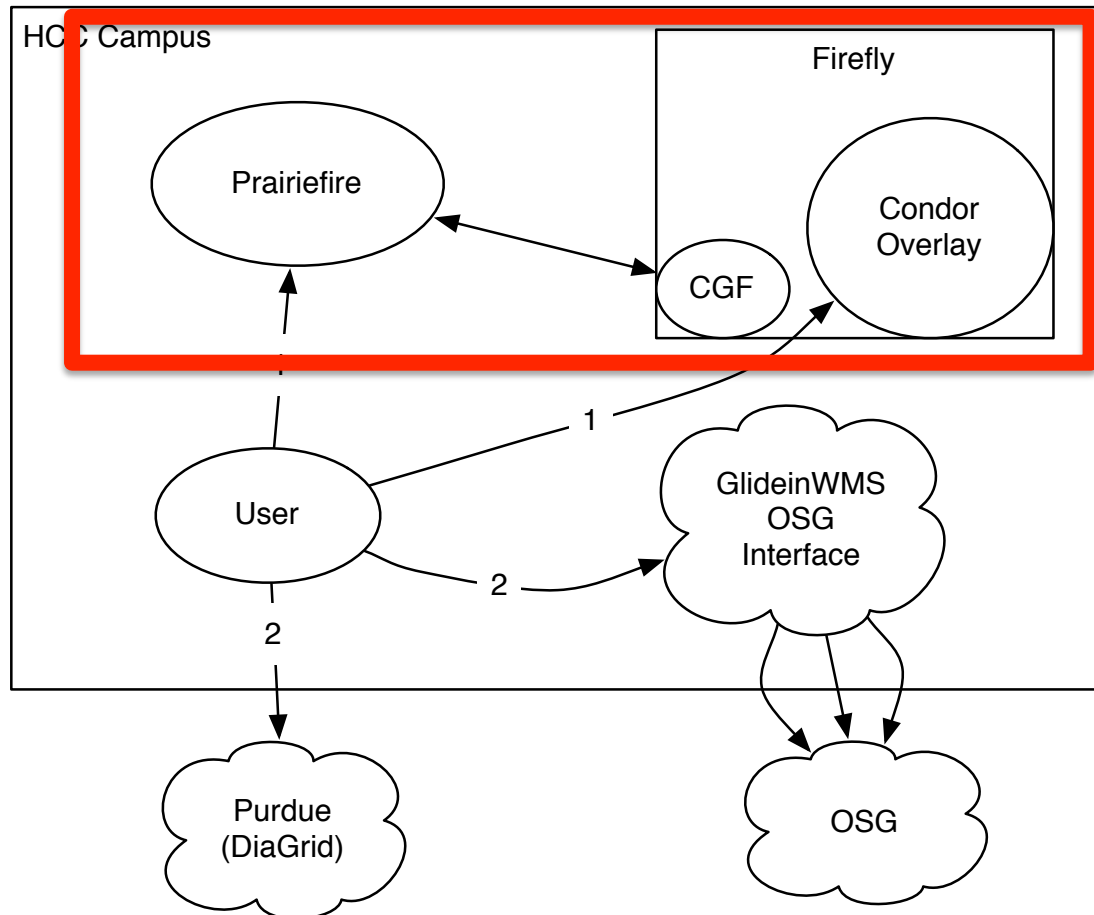# HCC's usage of GlideinWMS

- Center of our Campus grid
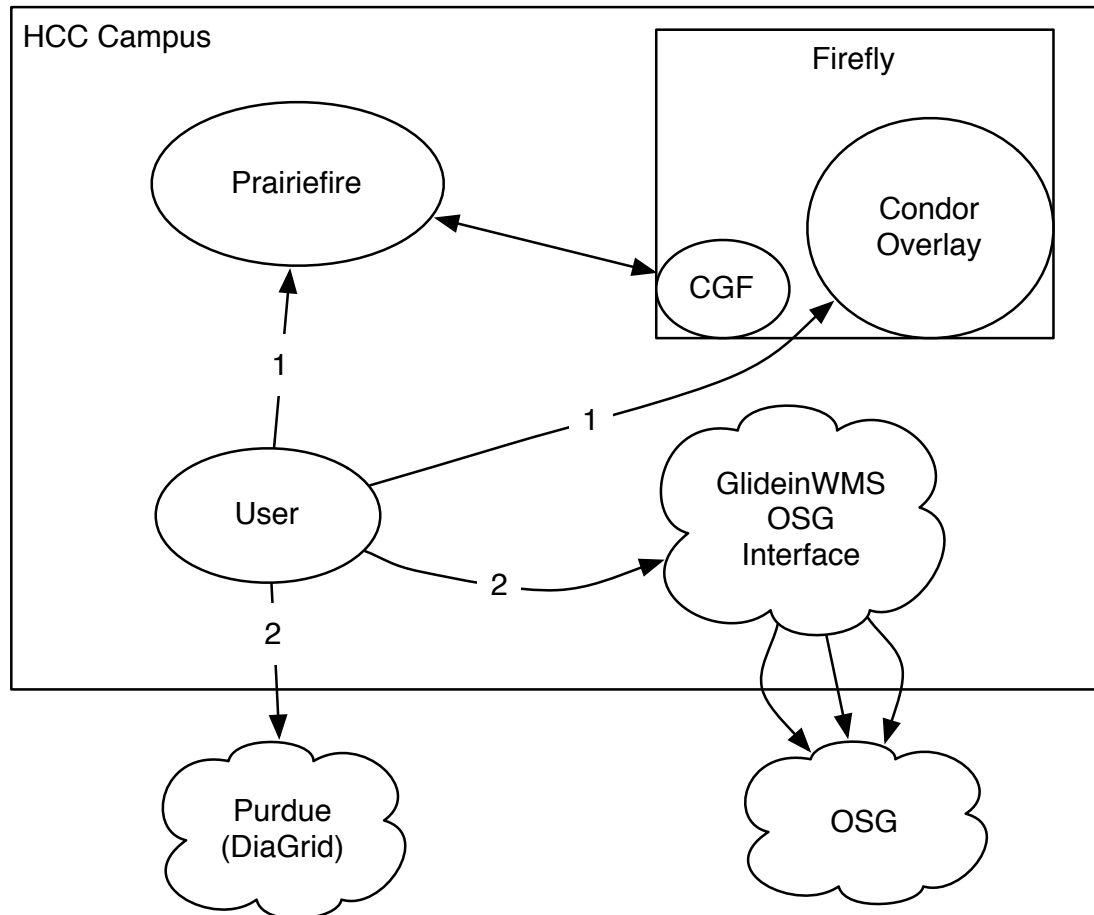
HCC's usage of GlideinWMS

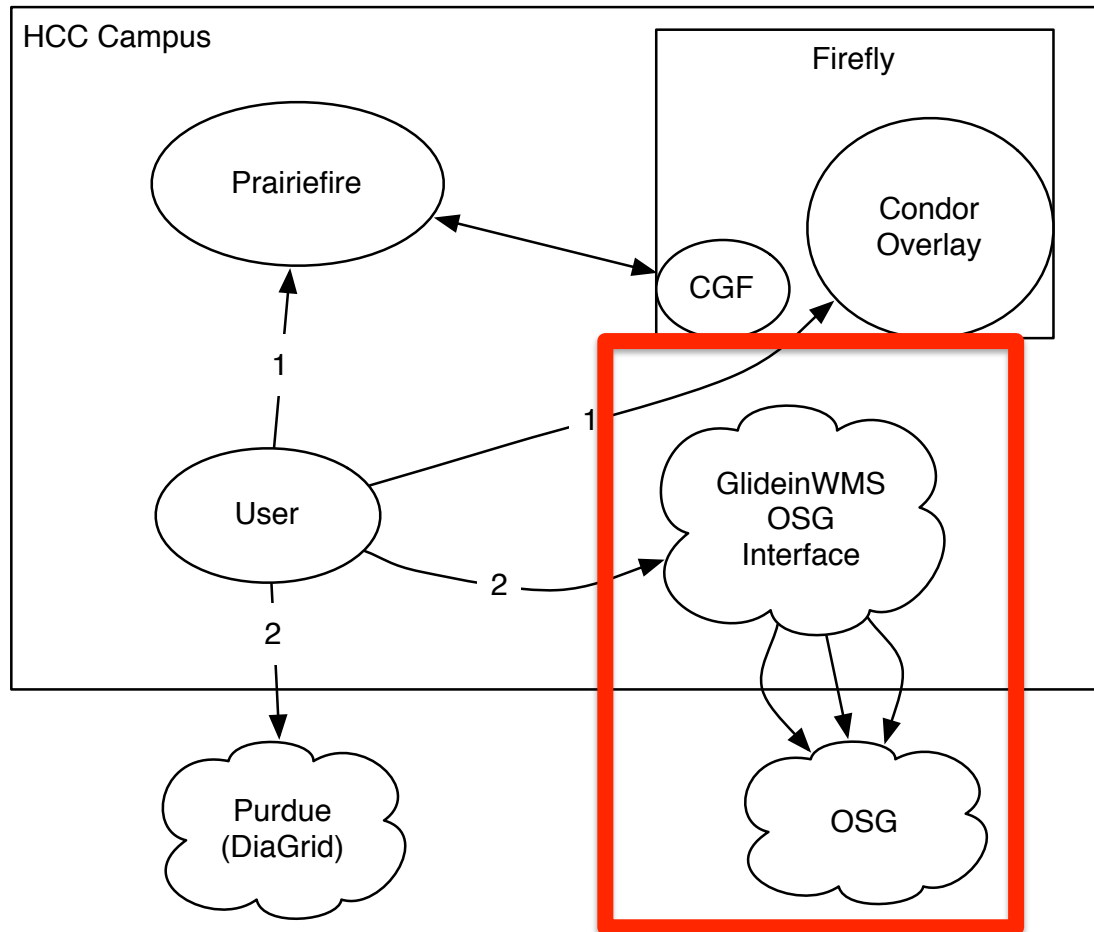- Users submit jobs on our glideinwms frontend

HCC's usage of GlideinWMS

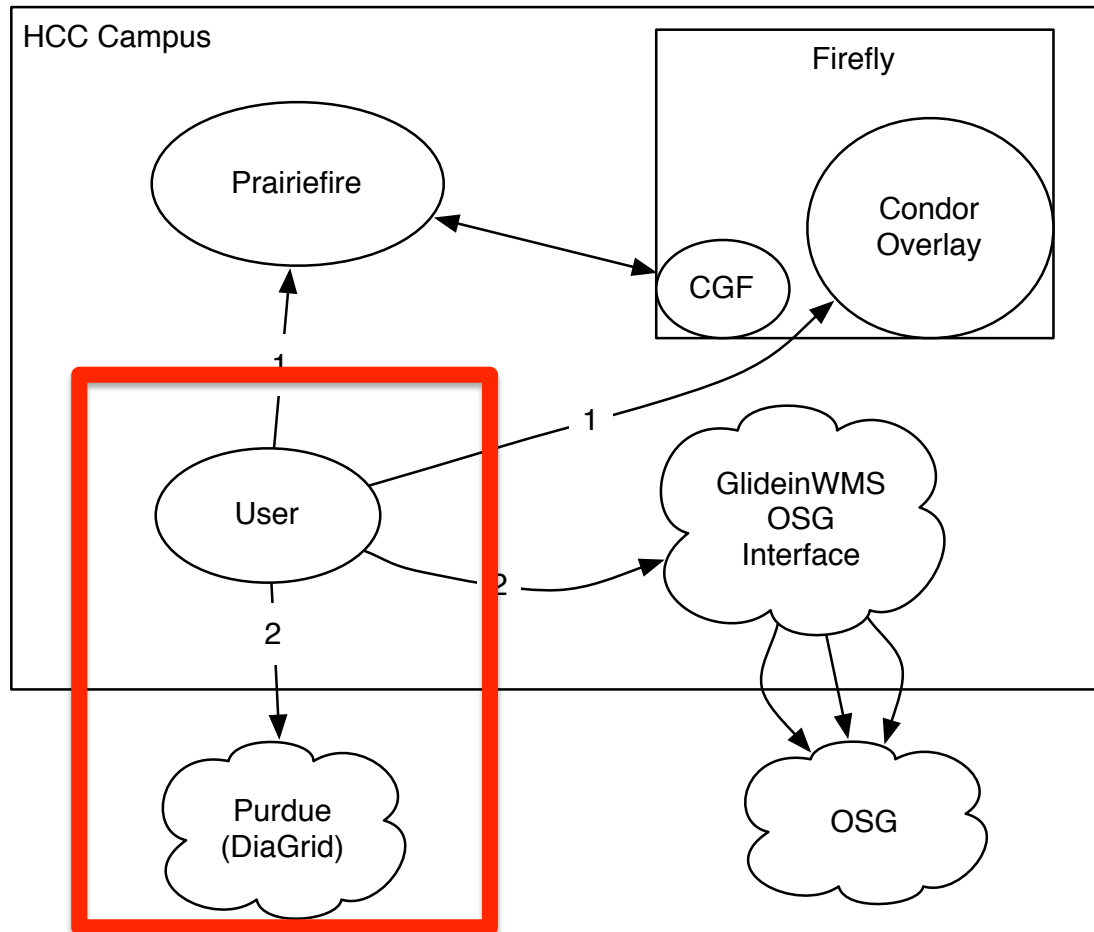- Jobs flock to on campus resources

HCC's usage of GlideinWMS

- Jobs go out to the OSG through GlideinWMS

HCC's usage of GlideinWMS

- Jobs flock out of our campus.

HCC's usage of GlideinWMS

- Developed custom monitoring

https://github.com/djw8605/GlideinMonitoring
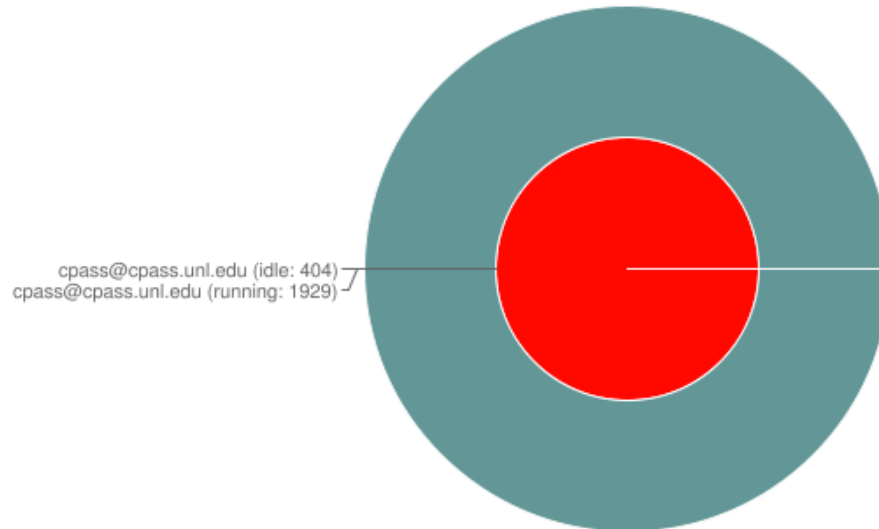
HCC's usage of GlideinWMS

- Contributed monitoring to GlideinWMS
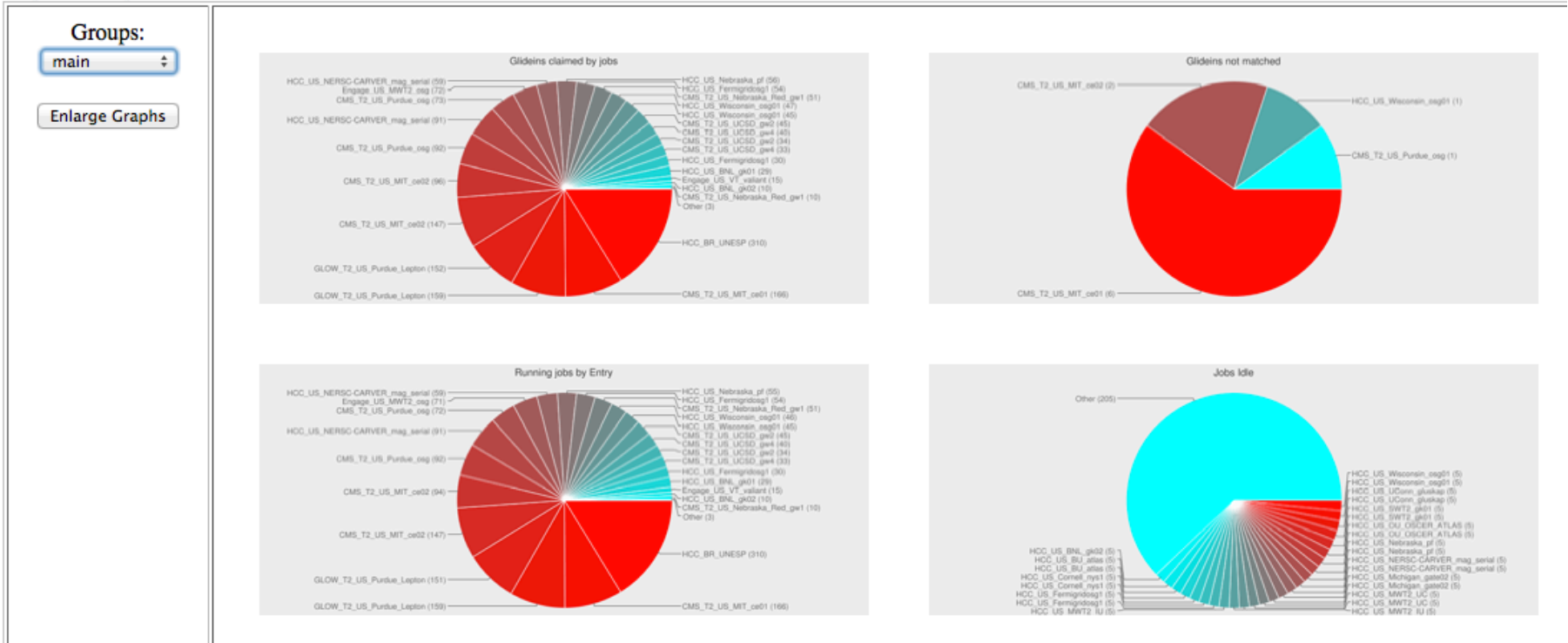
GlideinWMS Accounting

- Accounting is very important to HCC

- All of our accounting data is in a local Gratia Collector

- We created a gratia probe for GlideinWMS.
    - https://twiki.grid.iu.edu/bin/view/Accounting/ProbeConfigGlideinWMS

GlideinWMS Accounting

- See usage by user…

Daily Hours By User (Glidein)
14 Days from 2012-01-03 to 2012-01-17

Legend:
- Douglas Ryan Berry
- Tobias Toll
- Lana Muniz
- Yanjun Tu
- Linlin Zhang
- Michele Pioppi
- Wee Don Teo
- Ian Stokes-rees (portal)
- Other
- Lukasz Kreczko
- Vishagan Ratnaswamy
- Nicholas S Eggert
- Bruno Casal Larana
- Andreas Prlic
- Guangyou Qin
- Andrew Malone Melo
- Verena Ingrid Martinez Outschoorn
- Maurizio Pierini
- Troy Daniel Mulholland
- Brian R Drell

Maximum: 508,232 Hours, Minimum: 65,218 Hours, Average: 226,664 Hours, Current: 170,355 Hours

http://gratiaweb.grid.iu.edu/gratia/xml/glidein_dn_hours_bar

GlideinWMS Accounting

- And by site...

WMS Hours Spent on Jobs By Facility (Glidein)
14 Days from 2012-01-03 to 2012-01-17

Legend:
- Other
- GLOW
- Nebraska
- GRIF
- UKI-LT2-IC-HEP-overflow
- Purdue-RCAC
- Purdue-Rossmann
- FNAL_FERMIGRID
- INFN-PISA
- prairiefire
- UCSDT2
- CIT_CMS_T2
- UFlorida-HPC
- UFlorida-HPC-overflow
- UColorado_HEP-overflow
- USCMS-FNAL-WC1-CE
- GLOW-overflow
- UCSDT2-overflow
- UKI-SOUTHGRID-RALPP-overflow
- Firefly
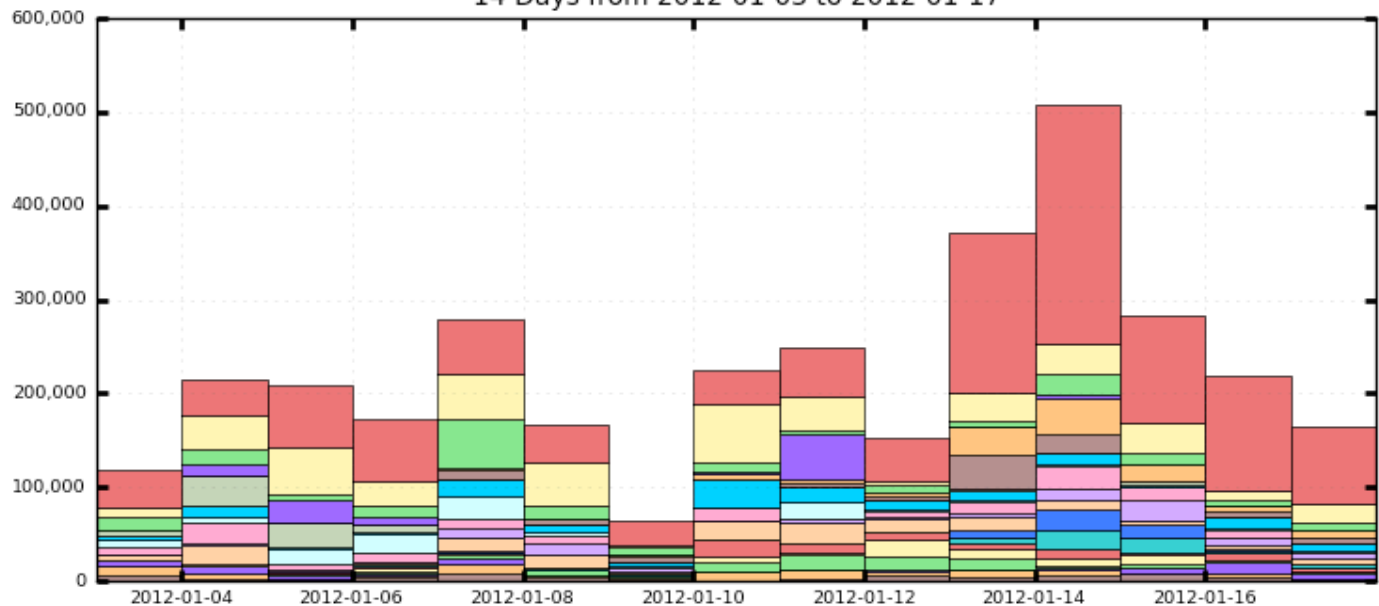
Maximum: 508,232 , Minimum: 65,129 , Average: 226,215 , Current: 163,718

http://gratiaweb.grid.iu.edu/gratia/xml/glidein_hours_bar_smry

GlideinWMS Accounting

- And get data in CSV format

```
Facility,Unknown Grouping,
IN2P3-IRES,1326758400,756.327777778
IN2P3-IRES,1326672000,220.241944444
IN2P3-IRES,1326585600,1991.90583333
IN2P3-IRES,1326499200,5364.97333333
IN2P3-IRES,1326412800,11331.8644444
IN2P3-IRES,1326326400,2801.0875
IN2P3-IRES,1326240000,2486.20472222
IN2P3-IRES,1326153600,2826.03833333
IN2P3-IRES,1326067200,1031.90277778
IN2P3-IRES,1325980800,112.641388889
IN2P3-IRES,1325894400,141.783611111
TTU-ANTAEUS,1326067200,0.169722222222
TTU-ANTAEUS,1325980800,280.915555556
TTU-ANTAEUS,1325894400,2587.67944444
TTU-ANTAEUS,1325808000,1009.18
TTU-ANTAEUS,1325721600,2505.24805556
TTU-ANTAEUS,1325635200,2446.05388889
TTU-ANTAEUS,1325548800,582.409166667
DESY-HH-overflow,1326758400,166.870555556
DESY-HH-overflow,1326672000,74.6447222222
DESY-HH-overflow,1326585600,41.3636111111
DESY-HH-overflow,1326499200,3530.63583333
DESY-HH-overflow,1326412800,7028.25472222
DESY-HH-overflow,1326326400,145.727777778
DESY-HH-overflow,1326240000,24.3452777778
DESY-HH-overflow,1326153600,754.48
```

http://gratiaweb-itb.grid.iu.edu/gratia/csv/glidein_hours_bar_smry

Web Portal to GlideinWMS

- Built bioinformatics (cpass) website submitting to GlideinWMS.

http://cpass.unl.edu/

Groups for Different Usage

- We have many glidein 'groups' for users
  - Main: most commonly used by generic submitters and cpass.unl.edu
  - Long: < 36 hour jobs
  - Extralong: < 72 hour jobs
  - HTPC: High throughput parallel computing jobs
 - T2Overflow: Transparent overflow for the Nebraska T2 to other sites.

Groups for Different Usage

- Requesting group is as simple as adding an attribute to job description

- For example, for a job running between ~16 (main max) and 36 hours, you could add:

```
universe = vanilla
…
+LongJob = TRUE
Requirements = GLIDEIN_Is_Long =?= TRUE
…
queue
```

Overflow Tier 2 through GlideWMS

- CMS jobs transparently flock to glideinwms frontend.

- Jobs are executed at other sites.
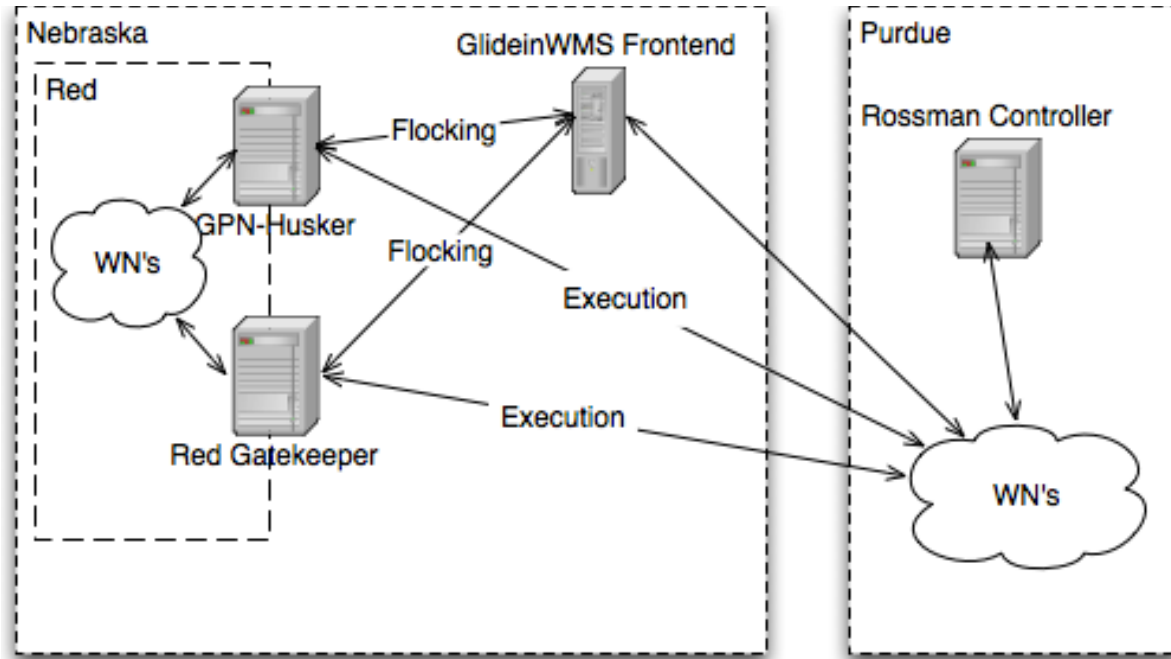
Comments on GlideinWMS

- Good
  - Huge improvement over Condor-G
  - Can setup and not touch for years.
- Bad
  - Complicated configuration
  - Resource utilization on frontend
  - Lack of ability for frontend operators to debug factory (pilot) issues.

Comments on OSG Glidein Factory

- Good
  - Responsive to questions and issues
  - Dual factories provide resiliency to bleeding edge UCSD
- Bad
  - Heard throttling at the factory is in the future (allocations?)